



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Model Reduction by Moment Matching for Linear Switched Systems

Bastug, Mert; Petreczky, Mihaly; Wisniewski, Rafal; Leth, John-Josef

Published in:
American Control Conference (ACC), 2014

DOI (link to publication from Publisher):
[10.1109/ACC.2014.6858983](https://doi.org/10.1109/ACC.2014.6858983)

Publication date:
2014

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Bastug, M., Petreczky, M., Wisniewski, R., & Leth, J.-J. (2014). Model Reduction by Moment Matching for Linear Switched Systems. In *American Control Conference (ACC), 2014* (pp. 3942 - 3947). IEEE Press. American Control Conference <https://doi.org/10.1109/ACC.2014.6858983>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Model Reduction by Moment Matching for Linear Switched Systems

Mert Baştuğ^{1,2}, Mihály Petreczky², Rafael Wisniewski¹ and John Leth¹

Abstract—A moment-matching method for the model reduction of linear switched systems (LSSs) is developed. The method is based upon a partial realization theory of LSSs and it is similar to the Krylov subspace methods used for moment matching for linear systems. The results are illustrated by a numerical example.

I. INTRODUCTION

A linear switched system (abbreviated by LSS) is a system which switches among a finite number of linear subsystems. That is, the state and output trajectory of an LSS is a concatenation of the state and output trajectories of the linear subsystems. At each time instance, the active linear subsystem is determined by a switching signal, which is considered as an additional external input to the system, i.e., any switching sequence is admissible. Linear switched systems represent a class of hybrid systems and they have been studied extensively, see [12], [25] for an overview.

Model reduction is the problem of approximating a dynamical system with another one of smaller complexity. For LSSs, by complexity, we will mean the number of continuous states, and by model reduction we will mean the approximation of the original LSS by another one, with a smaller number of continuous states.

Contribution of the paper In this paper, we present a model reduction algorithm based on the partial realization theory for LSSs [21]. The main idea is to replace the original LSS by an LSS of smaller order, such that a finite number of Markov parameters of the two LSSs coincide. By Markov parameters of the LSS we mean the Markov parameters of its input-output map, as defined in [21]. When applied to the linear case, the definition of [21] yields the usual definition of Markov parameters. The Markov parameters can be interpreted as high-order partial derivatives of the input-output maps with respect to the switching times. Hence, if those Markov parameters of the two LSSs which correspond to lower order partial derivatives coincide, then, by the same logic as used in Taylor series approximation, one could say that the two LSSs (more precisely, their input-output maps) are close. Hence, the proposed algorithm extends the well-known moment matching approach for linear systems [1]. By analogy with the linear case, we will refer to the proposed model reduction approach as *moment matching* for LSSs. The contribution of the paper is thus twofold: (1) the paper

formulates model reduction by moment matching for LSSs, and (2) it presents an algorithm for computing the reduced order system. The algorithm presented in this paper is similar to Krylov subspace methods for approximation of linear systems.

Motivation The order of the controller and the computation complexity of controller synthesis usually increase as the number of continuous states of the plant model increases. Hence, the smaller the plant model is, the easier it is to synthesize the control law and to implement it. This becomes especially relevant for hybrid systems, as many of the existing control synthesis methods are computationally demanding and result in large scale controllers. For example, many of the existing control synthesis methods rely on computing a finite-state abstraction of the plant model [26]. Often, the computational complexity of these methods and the size of the controller are exponential in the number of continuous states of the plant. This gets even worse for control problems with partial observation [18]. This means that even plant models of moderate size can become intractable and even a small reduction in the number of states can make a difference. For this reason, we expect that model reduction of switched systems will be useful for control of switched systems.

Related work In the linear case, model reduction is a mature research area, see [1] and the references therein. The subject of model reduction for hybrid and switched systems was addressed in several papers [4], [30], [15], [5], [9], [28], [29], [7], [10], [11], [16], [24]. Except [9], the cited papers propose techniques which involve solving certain LMIs, and for this reason, they tend to be applicable only to switched systems for which the continuous subsystems are stable. In contrast, the approach of this paper works for systems which are unstable. However, this comes at a price, since we are not able to propose analytic error bounds, like the ones for balanced truncation [23]. From a practical point of view, the lack of an analytic error bound need not be a very serious disadvantage, since it is often acceptable to evaluate the accuracy of the the approximation after the reduced model has been computed. For example, one can compute the L_2 distance between the original and reduced order model [23].

Note that the proposed algorithm does have a system theoretic interpretation, as it operates on the Markov parameters. Note also that the Markov parameters of switched systems characterize their input-output behavior uniquely [17]. Moreover, the Euclidean distance between Markov parameters can be used to define a natural distance for LSSs with state-space representations [22], [20]. For this reason, we believe that it might be possible to give a theoretical justification

¹Department of Electronic Systems, Automation and Control, Aalborg University, 9220 Aalborg, Denmark mertb@es.aau.dk, raf@es.aau.dk, jjl@es.aau.dk

²Department of Computer Science and Automatic Control (UR Informatique et Automatique), École des Mines de Douai, 59508 Douai, France mihaly.petreczky@mines-douai.fr

for the proposed algorithm. However, this remains a topic of future research. Note that even in the linear case, it is very challenging to give analytic error bounds for algorithms based on moment matching, [1].

The model reduction algorithm proposed in this paper is similar in spirit to moment matching for linear systems [1], [8] and bilinear systems [13], [2], [6]; however, the details and the system class considered are entirely different. The model reduction algorithm for LPV systems described in [27] is related, as it also relies on a realization algorithm and Markov parameters. In turn, the realization algorithms and Markov parameters of LPV systems and LSSs are closely related, [19]. However, the algorithm of [27] applies to a different system class, and it is not yet clear if it yields a partial realization.

Outline In Section II, we fix the notation and terminology of the paper. In Section III, we present the formal definition and main properties of LSSs. In Section IV, we recall the concept of Markov parameters, and we present the fundamental theorem and corollaries which form the basis of the model reduction by moment matching procedure. The algorithm itself is stated in Section V in detail. Finally, in Section VI the algorithm is illustrated on a numerical example.

II. PRELIMINARIES: NOTATION AND TERMINOLOGY

Denote by \mathbb{N} the set of natural numbers including 0. Denote by \mathbb{R}_+ the set $[0, +\infty)$ of nonnegative real numbers. In the sequel, let $PC(\mathbb{R}_+, S)$, with S a topological subspace of an Euclidean space \mathbb{R}^n , denote the set of *piecewise-continuous and left-continuous maps*. That is, $f \in PC(\mathbb{R}_+, S)$ if it has finitely many points of discontinuity on any compact subinterval of \mathbb{R}_+ , and at any point of discontinuity both the left-hand and right-hand side limits exists, and f is continuous from the left. In addition, denote by $AC(\mathbb{R}_+, \mathbb{R}^n)$ the set of *absolutely continuous maps*, and $L_{loc}(\mathbb{R}_+, \mathbb{R}^n)$ the set of *Lebesgue measurable maps* which are integrable on any compact interval.

III. LINEAR SWITCHED SYSTEMS

In this section, we present the formal definition of linear switched systems and recall a number of relevant definitions. We follow the presentation of [17], [23].

Definition 1 (LSS): A continuous time linear switched system (LSS) is a control system of the form

$$\frac{d}{dt}x(t) = A_{\sigma(t)}x(t) + B_{\sigma(t)}u(t), \quad x(t_0) = x_0 \quad (1a)$$

$$y(t) = C_{\sigma(t)}x(t) \quad (1b)$$

where¹ $Q = \{1, \dots, D\}$, $D > 0$, called the set of discrete modes, $\sigma \in PC(\mathbb{R}_+, Q)$ is called the switching signal, $u \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m)$ is called the input, $x \in AC(\mathbb{R}_+, \mathbb{R}^n)$ is called the state, and $y \in PC(\mathbb{R}_+, \mathbb{R}^p)$ is called the output. Moreover, $A_q \in \mathbb{R}^{n \times n}$, $B_q \in \mathbb{R}^{n \times m}$, $C_q \in \mathbb{R}^{p \times n}$ are the matrices of the

linear system in mode $q \in Q$, and x_0 is the initial state. The notation

$$\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) | q \in Q\}, x_0) \quad (2)$$

is used as a short-hand representation for LSSs of the form (1). The number n is called the *dimension (order)* of Σ and will sometimes be denoted by $\dim \Sigma$.

Throughout the paper, Σ denotes an LSS of the form (1). Next, we present the basic system theoretic concepts for LSSs.

Definition 2: The *input-to-state* map $X_{\Sigma, x}$ and *input-to-output* map $Y_{\Sigma, x}$ of Σ are the maps

$$X_{\Sigma, x} : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \rightarrow AC(\mathbb{R}_+, \mathbb{R}^n); (u, \sigma) \mapsto X_{\Sigma, x}(u, \sigma)$$

$$Y_{\Sigma, x} : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \rightarrow PC(\mathbb{R}_+, \mathbb{R}^p); (u, \sigma) \mapsto Y_{\Sigma, x}(u, \sigma)$$

defined by letting $t \mapsto X_{\Sigma, x}(u, \sigma)(t)$ be the solution to the Cauchy problem (1a) with $t_0 = 0$ and $x_0 = x$, and letting $Y_{\Sigma, x}(u, \sigma)(t) = C_{\sigma(t)}X_{\Sigma, x}(u, \sigma)(t)$ as in (1b).

The input-output behavior of an LSS realization can be formalized as a map

$$f : L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q) \rightarrow PC(\mathbb{R}_+, \mathbb{R}^p). \quad (3)$$

The value $f(u, \sigma)$ represents the output of the underlying (black-box) system. This system may or may not admit a description by an LSS. Next, we define when an LSS describes (realizes) a map of the form (3).

The LSS Σ of the form (1) is a *realization* of an input-output map f of the form (3), if f is the input-output map of Σ which corresponds to some initial state x_0 , i.e., $f = Y_{\Sigma, x_0}$. The map Y_{Σ, x_0} will be referred to as the *input-output map of* Σ , and it will be denoted by Y_Σ . The following discussion is only for realizable input-output maps.

We say that the LSSs Σ_1 and Σ_2 are *equivalent* if $Y_{\Sigma_1} = Y_{\Sigma_2}$. The LSS Σ_m is said to be a *minimal* realization of f , if Σ_m is a realization of f , and for any LSS Σ such that Σ is a realization of f , $\dim \Sigma_m \leq \dim \Sigma$. In [17], it is stated that an LSS realization Σ is minimal if and only if it is span-reachable and observable. See [17] for detailed definitions of span-reachability and observability for LSSs. Moreover, if Σ is a realization of f , then there exists an algorithm for computing from Σ a minimal realization Σ_m of f , see [17], [23]. Hence, in the sequel, unless stated otherwise we will assume that the LSSs are minimal realizations of their input-output maps.

IV. MOMENT MATCHING FOR LINEAR SWITCHED SYSTEMS: PROBLEM FORMULATION

In this section, we state formally the problem of moment matching for LSSs.

Notation 1: Consider a finite non-empty set Q with D elements, which will be called the *alphabet*. Denote by Q^* the set of finite sequences of elements of Q . The elements of Q^* are called *strings* or *words* over Q . Each non-empty word w is of the form $w = q_1 q_2 \dots q_k$ for some $q_1, q_2, \dots, q_k \in Q$. The element q_i is called the *i th letter* of w , for $i = 1, 2, \dots, k$ and k is called the *length* of w . The *empty sequence (word)* is denoted by ε . The length of word w is denoted by $|w|$;

¹ Q is endowed with discrete topology.

note that $|\varepsilon| = 0$. The set of non-empty words is denoted by Q^+ , i.e., $Q^+ = Q^* \setminus \{\varepsilon\}$. The *concatenation* of word $w \in Q^*$ with $v \in Q^*$ is denoted by wv : if $v = v_1 v_2 \cdots v_k$, and $w = w_1 w_2 \cdots w_m$, $k > 0, m > 0$, $v_1, v_2, \dots, v_k, w_1, w_2, \dots, w_m \in Q$, then $wv = v_1 v_2 \cdots v_k w_1 w_2 \cdots w_m$. If $v = \varepsilon$, then $wv = w$; if $w = \varepsilon$, then $wv = v$. For simplicity, the finite set Q will be identified with its index set, that is $Q = \{1, 2, \dots, D\}$.

Next, consider an input-output map f of the form (3). Notice that the restriction to a finite interval $[0, t]$ of any $\sigma \in PC(\mathbb{R}_+, Q)$ can be interpreted as finite sequence of elements of $Q \times \mathbb{R}_+$ of the form

$$\mu(\sigma) = (q_1, t_1)(q_2, t_2) \cdots (q_k, t_k) \quad (4)$$

where $q_1, \dots, q_k \in Q$ and $t_1, \dots, t_k \in \mathbb{R}_+ \setminus \{0\}$, $t_1 + \dots + t_k = t$, such that for all $s \in [0, t]$

$$\sigma(s) = \begin{cases} q_1 & \text{if } s \in [0, t_1] \\ q_2 & \text{if } s \in (t_1, t_1 + t_2] \\ \vdots & \\ q_i & \text{if } s \in (t_1 + \dots + t_{i-1}, t_1 + \dots + t_{i-1} + t_i] \\ \vdots & \\ q_k & \text{if } s \in (t_1 + \dots + t_{k-1}, t_1 + \dots + t_{k-1} + t_k] \end{cases}$$

From [17], it follows that a necessary condition for f to be realizable by an LSS is that f has a *generalized kernel representation*. For a detailed definition of a generalized kernel representation of f , we refer the reader to [17, Definition 19]. For our purposes, it is sufficient to recall that if f has a generalized kernel representation, then there exists a unique family of analytic functions $K_{q_1, \dots, q_k}^f : \mathbb{R}_+^k \rightarrow \mathbb{R}^p$ and $G_{q_1, \dots, q_k}^f : \mathbb{R}_+^k \rightarrow \mathbb{R}^{p \times m}$, $q_1, \dots, q_k \in Q$, $k \geq 1$, such that for all $(u, \sigma) \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q)$, $t > 0$

$$f(u, \sigma)(t) = K_{q_1 q_2 \dots q_k}^f(t_1, t_2, \dots, t_k) + \sum_{i=1}^k \int_0^{t_i} G_{q_i q_{i+1} \dots q_k}^f(t_i - s, t_{i+1}, \dots, t_k) u \left(s + \sum_{j=1}^{i-1} t_j \right) ds,$$

with $\mu(\sigma) = (q_1, t_1)(q_2, t_2) \cdots (q_k, t_k)$ and the functions $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$ satisfy a number of technical conditions, see [17, Definition 19] for details. From [17], it follows that there is a one-to-one correspondence between f and the family of maps $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$. If f has a realization by an LSS (1), then the functions $K_{q_1 q_2 \dots q_k}^f(t_1, t_2, \dots, t_k)$ and $G_{q_1 q_2 \dots q_k}^f(t_1, t_2, \dots, t_k)$ satisfy

$$K_{q_1 q_2 \dots q_k}^f(t_1, t_2, \dots, t_k) = C_{q_k} e^{A_{q_k} t_k} e^{A_{q_{k-1}} t_{k-1}} \cdots e^{A_{q_1} t_1} x_0$$

$$G_{q_1 q_2 \dots q_k}^f(t_1, t_2, \dots, t_k) = C_{q_k} e^{A_{q_k} t_k} e^{A_{q_{k-1}} t_{k-1}} \cdots e^{A_{q_1} t_1} B_{q_1}.$$

We can now define the Markov parameters of f as follows.

Definition 3 (Markov parameters): Let f have a GCR. The Markov parameters of f are the values of the map

$$M^f : Q^* \rightarrow \mathbb{R}^{Dp \times (mD+1)},$$

defined by,

$$M^f(v) = \begin{bmatrix} S_0(v1), & S(1v1), & \cdots & S(Dv1) \\ S_0(v2), & S(1v2), & \cdots & S(Dv2) \\ \vdots & \vdots & \cdots & \vdots \\ S_0(vD), & S(1vD), & \cdots & S(DvD) \end{bmatrix},$$

where the vectors $S_0(vq) \in \mathbb{R}^p$ and the matrices $S(q_0 vq) \in \mathbb{R}^{p \times m}$ are defined as follows. For all $q_0, q \in Q$,

$$S_0(q) = K_q^f(0) \text{ and } S(q_0 q) = G_{q_0 q}^f(0, 0).$$

and for all $q_0, q \in Q$, $v \in Q^*$, $v \neq \varepsilon$ by

$$S_0(vq) = \left. \frac{\partial}{\partial t_1} \cdots \frac{\partial}{\partial t_k} K_{q_1 \dots q_k}^f(t_1, \dots, t_k, 0) \right|_{t_1=t_2=\dots=t_k=0}$$

$$S(q_0 vq) = \left. \frac{\partial}{\partial t_1} \cdots \frac{\partial}{\partial t_k} G_{q_0 q_1 \dots q_k}^f(0, t_1, \dots, t_k, 0) \right|_{t_1=t_2=\dots=t_k=0}.$$

where $v = q_1 q_2 \cdots q_k$, $k \geq 0$, $q_1, q_2, \dots, q_k \in Q$.

That is, the Markov parameters of f are certain partial derivatives of the functions $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$. From [17], it follows that the Markov parameters $\{M^f(v)\}_{v \in Q^*}$ determine the maps $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$, and hence determine f , uniquely. In fact, in [17], it was shown that the entries of $\{M^f(v)\}_{v \in Q^*}$ can be considered as high-order derivatives of the maps $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$. More precisely, for any $\alpha_1, \dots, \alpha_k \in \mathbb{N}$, $q_1, \dots, q_k \in Q$, $k > 0$,

$$S_0(q_1^{\alpha_1} \cdots q_k^{\alpha_k+1}) = \left. \frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_k}}{\partial t_k^{\alpha_k}} K_{q_1 \dots q_k}^f(t_1, \dots, t_k) \right|_{t_1=\dots=t_k=0}$$

$$S(q_1^{\alpha_1+1} \cdots q_k^{\alpha_k+1}) = \left. \frac{\partial^{\alpha_1}}{\partial t_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_k}}{\partial t_k^{\alpha_k}} G_{q_1 \dots q_k}^f(t_1, \dots, t_k) \right|_{t_1=\dots=t_k=0}.$$

If f has a realization by an LSS Σ of the form (1), then the Markov-parameters of f can be expressed as products of the matrices of Σ . In order to present the corresponding formula, we will use the following notation.

Notation 2: Let $w = q_1 q_2 \cdots q_k \in Q^*$, $q_1, \dots, q_k \in Q$, $k > 0$ and $A_{q_i} \in \mathbb{R}^{n \times n}$, $i = 1, \dots, k$. Then the matrix A_w is defined as

$$A_w = A_{q_k} A_{q_{k-1}} \cdots A_{q_1}. \quad (5)$$

If $w = \varepsilon$, then A_ε is the identity matrix.

From [17], it follows that an LSS (1) is a realization of the map f if and only if f has a generalized kernel representation and $\forall v \in Q^* : S_0(vq) = C_q A_v x_0$ and $S(q_0 vq) = C_q A_v B_{q_0}$, or, in more compact form

$$\forall v \in Q^* : M^f(v) = \tilde{C} A_v \begin{bmatrix} x_0 & B_1 & B_2 & \cdots & B_D \end{bmatrix} \quad (6)$$

with $\tilde{C} = [C_1^T \cdots C_D^T]^T$.

The main idea behind moment matching for LSSs is as follows: approximate f by another input-output map \hat{f} , such that some of the Markov parameters of f and \hat{f} coincide. One obvious choice is to say that $M^f(v) = M^{\hat{f}}(v)$ for all $v \in Q^*$, $|v| \leq N$ for some N . Intuitively, this means that all the partial derivatives of order at most N of $\{K_{q_1 \dots q_k}^f, G_{q_1 \dots q_k}^f \mid q_1, \dots, q_k \in Q, k \geq 0\}$ and of $\{K_{q_1 \dots q_k}^{\hat{f}}, G_{q_1 \dots q_k}^{\hat{f}} \mid q_1, \dots, q_k \in Q, k \geq 0\}$ coincide. Intuitively, this will mean that for any input and switching signal $(u, \sigma) \in L_{loc}(\mathbb{R}_+, \mathbb{R}^m) \times PC(\mathbb{R}_+, Q)$, the outputs $f(u, \sigma)(t)$ and $\hat{f}(u, \sigma)(t)$ should be close, at least for small enough t . Moreover, since we are interested in input-output maps which are realizable by LSSs, we would like \hat{f} to be realizable by an LSS of dimension smaller than

that of the minimal LSS realization f . In order to formalize the intuition described above, we introduce the following definitions.

Definition 4 (N-partial realization): The LSS (1) is called *N-partial realization* of f , if

$$\forall v \in Q^*, |v| \leq N : M^f(v) = \tilde{C}A_v\tilde{B}$$

with $\tilde{C} = [C_1^T \ \dots \ C_D^T]^T$ and $\tilde{B} = [x_0 \ B_1 \ B_2 \ \dots \ B_D]$. If Σ is of the form (1) and Y_Σ is the input-output map of Σ , then the concept of N partial realization can be interpreted as follows: Σ is an N partial realization of f , if those Markov parameters of f and Y_Σ which are indexed by words of length at most N coincide. The problem of model reduction by moment matching can now be formulated as follows.

Problem 1 (Moment matching): Let Σ be an LSS (1) and let $f = Y_\Sigma$ be its input-output map. Fix $N \in \mathbb{N}$. Find an LSS $\tilde{\Sigma}$ such that $\dim \tilde{\Sigma} < \dim \Sigma$ and $\tilde{\Sigma}$ is an N partial realization of $f = Y_\Sigma$.

Note that there is a trade off between the choice of N and the dimension Σ . This follows from the result of [21, Theorem 4].

Theorem 1: Assume that Σ is a minimal realization of f and $2\dim \Sigma - 1 \leq N$. Then for any LSS $\tilde{\Sigma}$ which is an N partial realization of f , $\dim \Sigma \leq \dim \tilde{\Sigma}$.

That is, if we choose N too high, namely if we choose any N such that $N \geq 2n - 1$, where n is the dimension of a minimal LSSs realization of f , then there will be no hope of finding an LSS which is an N partial realization of the original input-output map, and whose dimension is lower than n .

In order to solve the moment matching problem, one could consider applying the partial realization algorithm [21]. However, that approach yields a model reduction algorithm whose memory-usage and run-time complexity is exponential.

V. THE MODEL REDUCTION ALGORITHM

In this section, the aim is to present an efficient model reduction algorithm which transforms an LSS Σ into an LSS $\tilde{\Sigma}$ such that $\dim \tilde{\Sigma} \leq \dim \Sigma$ and $\tilde{\Sigma}$ is an N partial realization of the input-output map of Σ . The presented algorithm has polynomial computational complexity and does not involve the explicit computation of the Hankel matrix as in [21].

In the sequel, the image (column space) of a real matrix M is denoted by $\text{im}(M)$ and $\text{rank}(M)$ is the dimension of $\text{im}(M)$. We will start with presenting the following definitions.

Definition 5 ((Partial) Unobservability subspace): For an LSS Σ , and $N \in \mathbb{N}$ define the N step unobservability subspace as

$$\mathcal{O}_N(\Sigma) = \bigcap_{v \in Q^*, |v| \leq N, q \in Q} \ker(C_q A_v).$$

If Σ is clear from the context, we will denote $\mathcal{O}_N(\Sigma)$ by \mathcal{O}_N . It is not difficult to see that $\mathcal{O}_0 = \bigcap_{q \in Q} \ker(C_q)$ and for any $N > 0$, $\mathcal{O}_N = \mathcal{O}_0 \cap \bigcap_{q \in Q} \mathcal{O}_{N-1} A_q$. From [25], [17], it follows that Σ is observable if and only if $\mathcal{O}_N(\Sigma) = \{0\}$ for all $N \geq \dim \Sigma - 1$.

Definition 6 ((Partial) Reachability space): For an LSS Σ , define the N step reachability space as follows:

$$\mathcal{R}_N(\Sigma) = \text{Span}\{\text{im}(A_v \tilde{B}) \mid v \in Q^*, |v| \leq N\}, \quad (7)$$

where $\tilde{B} = [x_0, B_1, B_2, \dots, B_D]$. If Σ is clear from the context, we will denote $\mathcal{R}_N(\Sigma)$ by \mathcal{R}_N .

It is easy to see that $\mathcal{R}_0 = \text{im}(\tilde{B})$ and $\mathcal{R}_N = \text{im}(\tilde{B}) + \sum_{q \in Q} A_q \mathcal{R}_{N-1}$, for $N > 0$. It follows from [17], [25] that Σ is span-reachable if and only if $\dim \mathcal{R}_N = n$ for all $N \geq n - 1$.

Theorem 2 (One sided moment matching (columns)):

Let $\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ be an LSS realization of the input-output map f , $P \in \mathbb{R}^{n \times r}$ be a full column rank matrix such that

$$\mathcal{R}_N(\Sigma) = \text{im}(P).$$

If $\tilde{\Sigma} = (p, m, r, Q, \{(\tilde{A}_q, \tilde{B}_q, \tilde{C}_q) \mid q \in Q\}, \tilde{x}_0)$ is an LSS such that for each $q \in Q$, the matrices $\tilde{A}_q, \tilde{B}_q, \tilde{C}_q$ and the vector \tilde{x}_0 are defined as

$$\tilde{A}_q = P^{-1} A_q P, \tilde{B}_q = P^{-1} B_q, \tilde{C}_q = C_q P, \tilde{x}_0 = P^{-1} x_0,$$

where P^{-1} is a left inverse of P , then $\tilde{\Sigma}$ is an N -partial realization of f .

Proof: See [3]. ■

Using a dual argument, the following result can be proven.

Theorem 3 (One sided moment matching (observability)):

Let $\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ be an LSS realization of the input-output map f , $W \in \mathbb{R}^{r \times n}$ be a full row rank matrix such that

$$\mathcal{O}_N(\Sigma) = \ker(W)$$

Let W^{-1} be any right inverse of W and let $\tilde{\Sigma} = (p, m, r, Q, \{(\tilde{A}_q, \tilde{B}_q, \tilde{C}_q) \mid q \in Q\}, \tilde{x}_0)$ be an LSS such that for each $q \in Q$, the matrices $\tilde{A}_q, \tilde{B}_q, \tilde{C}_q$ and the vector \tilde{x}_0 are defined as

$$\tilde{A}_q = W A_q W^{-1}, \tilde{B}_q = W B_q, \tilde{C}_q = C_q W^{-1}, \tilde{x}_0 = W x_0.$$

Then $\tilde{\Sigma}$ is an N -partial realization of f .

Finally, by combining the proofs of Theorem 2 and Theorem 3, the following can be shown.

Theorem 4 (Two sided moment matching): Let

$\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) \mid q \in Q\}, x_0)$ be an LSS realization of the input-output map f , $V \in \mathbb{R}^{n \times r}$ and $W \in \mathbb{R}^{r \times n}$ be respectively full column rank and full row rank matrices such that

$$\mathcal{R}_N(\Sigma) = \text{im}(V), \mathcal{O}_N(\Sigma) = \ker(W) \text{ and } \text{rank}(WV) = r.$$

If $\tilde{\Sigma} = (p, m, r, Q, \{(\tilde{A}_q, \tilde{B}_q, \tilde{C}_q) \mid q \in Q\}, \tilde{x}_0)$ is an LSS such that for each $q \in Q$, the matrices $\tilde{A}_q, \tilde{B}_q, \tilde{C}_q$ and the vector \tilde{x}_0 are defined as

$$\tilde{A}_q = W A_q V (WV)^{-1}, \tilde{B}_q = W B_q, \tilde{C}_q = C_q V (WV)^{-1}, \tilde{x}_0 = W x_0,$$

then $\tilde{\Sigma}$ is a $2N$ -partial realization of f .

Now, we will present an efficient algorithm of model reduction by moment matching, which computes either an N or $2N$ -partial realization $\tilde{\Sigma}$ for an f which is realized by an LSS Σ . First, we present algorithms for computing the subspaces \mathcal{R}_N and \mathcal{O}_N . To this end, we will use the following notation: if M is any real matrix, then denote by **orth**(M) the matrix U such that U is full column rank, $\text{rank}(U) = \text{rank}(M)$, $\text{im}(U) = \text{im}(M)$ and $U^T U = I$. Note

that U can easily be computed from M numerically, see for example the Matlab command `orth`. The algorithm for computing \mathcal{R}_N is presented in Algorithm 1 below. By duality,

Algorithm 1 Calculate a matrix representation of \mathcal{R}_N ,

Inputs: $(\{A_q, B_q\}_{q \in Q}, x_0)$ and N

Outputs: $P \in \mathbb{R}^{n \times r}$ such that $\text{rank}(P) = r$, $\text{im}(P) = \mathcal{R}_N$.

```

 $P := U_0$ ,  $U_0 := \text{orth}[x_0, B_1, \dots, B_D]$ .
for  $k = 1 \dots N$  do
   $P := \text{orth}([U_0, A_1 P, A_2 P, \dots, A_D P])$ 
end for
return  $P$ .
```

we can use Algorithm 1 to compute \mathcal{O}_N , the details are presented in Algorithm 2.

Algorithm 2 Calculate a matrix representation of \mathcal{O}_N

Inputs: $\{A_q, C_q\}_{q \in Q}$ and N

Output: $W \in \mathbb{R}^{r \times n}$, such that $\text{rank}(W) = r$ and $\ker(W) = \mathcal{O}_N$.

```

Apply Algorithm 1 with inputs  $(\{A_q^T, C_q^T\}_{q \in Q}, 0)$  to obtain
a matrix  $P$ .
return  $W = P^T$ .
```

Notice that the computational complexity of Algorithm 1 and Algorithm 2 is polynomial in N and n , even though the spaces of \mathcal{R}_N (resp. \mathcal{O}_N) are generated by images (resp. kernels) of exponentially many matrices.

Using Algorithm 1 and 2, we can formulate a model reduction algorithm, see Algorithm 3.

Algorithm 3 Moment matching for LSSs

Inputs: $\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) | q \in Q\}, x_0)$ and $N \in \mathbb{N}$.

Output: $\tilde{\Sigma} = (p, m, r, Q, \{(\bar{A}_q, \bar{B}_q, \bar{C}_q) | q \in Q\}, \bar{x}_0)$.

```

Using Algorithm 1-2 compute matrices  $P$  and  $W$  such that
 $P$  is full column rank,  $W$  is full row rank and  $\text{im}(P) = \mathcal{R}_N$ ,
 $\ker(W) = \mathcal{O}_N$ .
```

```

if  $\text{rank}(P) = \text{rank}(W) = \text{rank}(WP)$  then
```

```

  Let  $r = \text{rank}(P)$  and
```

$$\bar{A}_q = WA_q P(WP)^{-1}, \bar{C}_q = C_q P(WP)^{-1},$$

$$\bar{B}_q = WB_q, \bar{x}_0 = Wx_0.$$

```

end if
```

```

if  $\text{rank}(P) \geq \text{rank}(W)$  then
```

```

  Let  $r = \text{rank}(P)$ ,  $P^{-1}$  be a left inverse of  $P$  and set
```

$$\bar{A}_q = P^{-1} A_q P, \bar{C}_q = C_q P, \bar{B}_q = P^{-1} B_q, \bar{x}_0 = P^{-1} x_0.$$

```

end if
```

```

if  $\text{rank}(P) < \text{rank}(W)$  then
```

```

  Let  $r = \text{rank}(W)$  and let  $W^{-1}$  be a right inverse of  $W$ .
  Set
```

$$\bar{A}_q = WA_q W^{-1}, \bar{C}_q = C_q W^{-1}, \bar{B}_q = WB_q, \bar{x}_0 = Wx_0.$$

```

end if
```

```

return  $\tilde{\Sigma} = (p, m, r, Q, \{(\bar{A}_q, \bar{B}_q, \bar{C}_q) | q \in Q\}, \bar{x}_0)$ .
```

Theorem 2 – 4 imply the following corollary on correctness of Algorithm 3.

Corollary 1 (Correctness of Algorithm 3): Using the notation of Algorithm 3, the following holds: If $\text{rank}(P) = \text{rank}(W) = \text{rank}(WP)$, then Algorithm 3 returns an $2N$ partial realization of $f = Y_\Sigma$. Otherwise, Algorithm 3 returns an N partial realization of $f = Y_\Sigma$.

Remark 1 (Implementation): The implementation of Algorithm 3 in MATLAB is available from <https://kom.aau.dk/~mertb/>.

Remark 2 (Computational complexity): The memory and run time complexity of Algorithm 3 is polynomial in N, D and n . For details, see [3].

VI. NUMERICAL EXAMPLE

In this section, a generic numerical example is presented to illustrate the model reduction procedure. The procedure is applied to get a reduced order approximation to an LSS whose local modes are unstable. The original LSS used in this case is an LSS of the form $\Sigma = (p, m, n, Q, \{(A_q, B_q, C_q) | q \in Q\}, x_0)$ with $p = m = 1$, $n = 12$ and $Q = \{1, 2\}$. The resulting reduced order model $\tilde{\Sigma}_1$ is a 1-partial realization of Y_Σ of order 9. The data of A_q, B_q, C_q parameters and the initial state x_0 used for simulation are generated randomly and they are available from <https://kom.aau.dk/~mertb/>. A random switching signal with minimum dwell time (time between two subsequent changes in the switching signal) of 0.1 and a random input signal $u(t)$ with uniform distribution is used for simulation. The simulation time interval is $t = [0, 3]$. For $N = 1$, an approximation LSS of order 9 whose Markov parameters indexed by the words of length at most 1 are matched with the original LSS Σ is acquired, i.e the original LSS Σ is approximated by a 1-partial realization $\tilde{\Sigma}_1$. Note that the precise number of matched Markov parameters is equal to the number of words in the set Q^* of length at most $N = 1$, which is in this case 3. The output $y(t)$ of the original system Σ and the output $\bar{y}(t)$ of the reduced order system $\tilde{\Sigma}_1$ are simulated with the parameters given for 500 random switching sequences and input trajectories. For each simulation, the responses of the original and reduced order LSSs are compared with the best fit rate (BFR) (see [14], [27]) which is defined as

$$\text{BFR} = 100\% \max \left(1 - \frac{\|y(\cdot) - \bar{y}(\cdot)\|_2}{\|y(\cdot) - y_m\|_2}, 0 \right)$$

where y_m is the mean of y and $\|\cdot\|_2$ is the ℓ_2 norm. Even though the BFR is defined for output sequences rather than functions with the domain \mathbb{R}_+ , we could still apply it, since y and \bar{y} were obtained by computing a numerical solution of the LSS, and as a result they are both defined on a discretization of the time axis, i.e., y and \bar{y} are arrays containing the output values in the sampled time instances. For this example, the mean of the BFRs for 500 simulations is acquired as 79.0518%; whereas, the best acquired BFR is 90.8013% and the worst is 62.7846%. The outputs $y(t)$ and $\bar{y}(t)$ of the most successful simulation are shown in Fig. 1.

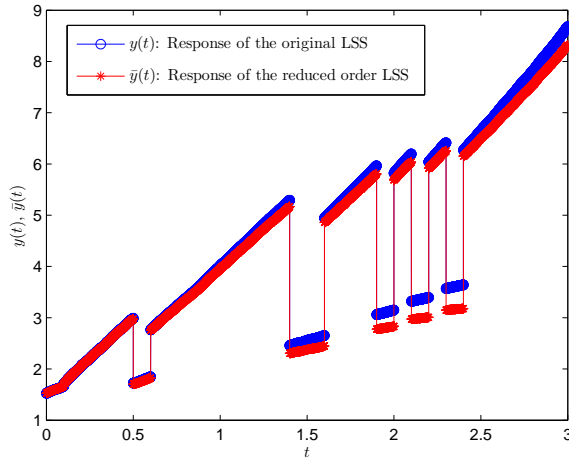


Fig. 1. The response $y(t)$ of the original LSS Σ of order 12 and the response $\bar{y}(t)$ of the reduced order approximation LSS $\tilde{\Sigma}_1$ of order 9.

VII. CONCLUSIONS

A moment matching procedure for model reduction of LSSs has been given. It has been proven that with this procedure, as long as a certain criterion is satisfied, it is possible to acquire at least one reduced order approximation to the original LSS whose first some number of Markov parameters are matched with the original one's. The procedure is based on constructing matrices whose image or kernel is the partial reachability or unobservability subspaces of an LSS respectively. Since we do not explicitly compute the Hankel matrices, the computational complexity does not increase exponentially with the number of moments to be matched, which is particularly important for large scale systems.

REFERENCES

- [1] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, Philadelphia, PA, 2005.
- [2] Z. Bai and D. Skoogh. A projection method for model reduction of bilinear dynamical systems. *Linear Algebra and its Applications*, 415(2 - 3):406 – 425, 2006.
- [3] M. Bastug, M. Petreczky, R. Wisniewski, and John Leth. Model reduction by moment matching for linear switched systems. Technical Report arXiv:1403.1734, ArXiv, 2014. Available at <http://arxiv.org/abs/1403.1734>.
- [4] A. Birouche, J. Guilet, B. Mourillon, and M. Basset. Gramian based approach to model order-reduction for discrete-time switched linear systems. In *Proc. Mediterranean Conference on Control and Automation*, 2010.
- [5] Y. Chahlaoui. Model reduction of hybrid switched systems. In *Proceeding of the 4th Conference on Trends in Applied Mathematics in Tunisia, Algeria and Morocco, May 4-8, Kenitra, Morocco*, 2009.
- [6] G. M. Flagg. *Interpolation Methods for the Model Reduction of Bilinear Systems*. PhD thesis, Virginia Polytechnic Institute, 2012.
- [7] H. Gao, J. Lam, and C. Wang. Model simplification for switched hybrid systems. *Systems & Control Letters*, 55:1015–1021, 2006.
- [8] S. Gugercin. *Projection methods for model reduction of large-scale dynamical systems*. PhD thesis, Rice Univ., Houston, TX, May 2003.
- [9] C.G.J.M. Habets and J. H. van Schuppen. Reduction of affine systems on polytopes. In *International Symposium on Mathematical Theory of Networks and Systems*, 2002.

- [10] G. Kotsalis, A. Megretski, and M. A. Dahleh. Balanced truncation for a class of stochastic jump linear systems and model reduction of hidden Markov models. *IEEE Transactions on Automatic Control*, 53(11), 2008.
- [11] G. Kotsalis and A. Rantzer. Balanced truncation for discrete-time Markov jump linear systems. *IEEE Transactions on Automatic Control*, 55(11), 2010.
- [12] D. Liberzon. *Switching in Systems and Control*. Birkhäuser, Boston, MA, 2003.
- [13] Y. Lin, L. Bao, and Y. Wei. A model-order reduction method based on krylov subspaces for mimo bilinear dynamical systems. *Journal of Applied Mathematics and Computing*, 25(1-2):293–304, 2007.
- [14] L. Ljung. *System Identification, Theory for the User*. Prentice Hall, Englewood Cliffs, NJ, 1999.
- [15] E. Mazzi, A.S. Vincentelli, A. Balluchi, and A. Bicchi. Hybrid system model reduction. In *IEEE International conference on Decision and Control*, 2008.
- [16] N. Monshizadeh, H. Trentelman, and M. Camlibel. A simultaneous balanced truncation approach to model reduction of switched linear systems. *Automatic Control, IEEE Transactions on*, PP(99):1, 2012.
- [17] M. Petreczky. Realization theory for linear and bilinear switched systems: formal power series approach - part i: realization theory of linear switched systems. *ESAIM Control, Optimization and Calculus of Variations*, 17:410–445, 2011.
- [18] M. Petreczky, P. Collins, D.A. van Beek, J.H. van Schuppen, and J.E. Rooda. Sampled-data control of hybrid systems with discrete inputs and outputs. In *Proceedings of 3rd IFAC Conference on Analysis and Design of Hybrid Systems (ADHS09)*, 2009.
- [19] M. Petreczky and G. Mercère. Affine LPV systems: realization theory, input-output equations and relationship with linear switched systems. In *Proceedings of the IEEE Conference on Decision and Control*, Maui, Hawaii, USA, December 2012.
- [20] M. Petreczky and R. Peeters. Spaces of nonlinear and hybrid systems representable by recognizable formal power series. In *Proc. 19th International Symposium on Mathematical Theory of Networks and Systems*, pages 1051–1058, Budapest, Hungary, July 2010.
- [21] M. Petreczky and J. H. van Schuppen. Partial-realization theory for linear switched systems - a formal power series approach. *Automatica*, 47:2177–2184, October 2011.
- [22] M. Petreczky and R. Vidal. Metrics and topology for nonlinear and hybrid systems. *Hybrid Systems: Computation and Control (HSCC) 2007, LNCS*, 4416:459–472, April 2007.
- [23] M. Petreczky, R. Wisniewski, and J. Leth. Balanced truncation for linear switched systems. *Nonlinear Analysis: Hybrid Systems*, 10:4–20, November 2013.
- [24] H.R. Shaker and R. Wisniewski. Generalized gramian framework for model/controller order reduction of switched systems. *International Journal of Systems Science, in press*, 2011.
- [25] Z. Sun and S. S. Ge. *Switched linear systems : control and design*. Springer, London, 2005.
- [26] P. Tabuada. *Verification and Control of Hybrid Systems: A Symbolic Approach*. Springer-Verlag, 2009.
- [27] R. Tóth, H. S. Abbas, and H. Werner. On the state-space realization of lpv input-output models: Practical approaches. *IEEE Trans. Contr. Syst. Technol.*, 20:139–153, January 2012.
- [28] L. Zhang, E. Boukas, and P. Shi. Mu-dependent model reduction for uncertain discrete-time switched linear systems with average dwell time. *International Journal of Control*, 82(2):378– 388, 2009.
- [29] L. Zhang and P. Shi. Model reduction for switched lpv systems with average dwell time. *IEEE Transactions on Automatic Control*, 53:2443–2448, 2008.
- [30] L. Zhang, P. Shi, E. Boukas, and C. Wang. Model reduction for uncertain switched linear discrete-time systems. *Automatica*, 44(11):2944 – 2949, 2008.